



# Analyzing Domain Knowledge for Big Data Analysis: A Case Study with Urban Tree Type Classification

Samantha Detor<sup>1,2</sup>, Abigail Roh<sup>1,2</sup>, Yiqun Xie<sup>2(✉)</sup>,  
and Shashi Shekhar<sup>2</sup>

<sup>1</sup> Breck School, Minneapolis, MN 55422, USA  
{detosa, rohab}@student.breckschool.org

<sup>2</sup> Department of Computer Science and Engineering, University of Minnesota,  
Minneapolis, MN 55455, USA  
{xiexx347, shekhar}@umn.edu

**Abstract.** The goals of this research were to create a labeled dataset of tree shadows and to test the feasibility of shadow-based tree type identification using aerial imagery. Urban tree big data that provides information about individual trees can help city planners optimize positive benefits of urban trees (e.g., increasing wellbeing of city residents) while managing potential negative impacts (e.g., risk to power lines). The continual rise of tree type specific threats, such as emerald ash borer, due to climate change has made this problem more pressing in recent years. However, urban tree big data are time consuming to create. This paper evaluates the potential of a new tree type identification method that utilizes shadows in aerial imagery to survey larger regions of land in a shorter amount of time. This work is challenging because there are structural variations across a given tree type and few verified tree type identification datasets exist. Related work has not explored how tree structure characteristics translate into a profile view of a tree's shadow or quantified the feasibility of shadow-only based tree type identification. We created a consistent and accurate dataset of 4,613 tree shadows using ground truthing procedures and novel methods for ensuring consistent collection of spatial shadow data that take binary and spatial agreement between raters into account. Our results show that identifying trees from shadows in aerial imagery is feasible and merits further exploration in the future.

**Keywords:** Aerial imagery · Urban trees · Labeled data creation · Ground truthing · Smart cities · Tree disease · Urban tree big data

## 1 Introduction

The goal of this study was to maximize the accuracy of classification of tree types based on corresponding shadows in aerial imagery. The study flow is depicted in Fig. 1. The image on the left displays a subsection of the high spatial resolution aerial

---

S. Detor and A. Roh—These authors contributed equally to this work.

imagery used as an input in this study. The image in the center shows the same region with its corresponding final shadow annotations, which were part of the dataset created and utilized in this research. The third box shows the output of this research, which was synthesized through the analysis of the dataset created in the intermediate step of this research.



**Fig. 1.** Study flow.

Often overlooked, trees are a key part of a sustainable, content, and safe city. Urban trees provide shade, clean air, and increase the wellbeing of city residents. A study in Chicago showed that trees help build community and lessen crime, and a view of trees from an office window increases job satisfaction and well-being [15, 22]. Information on the location of specific types of trees can assist city leaders in making sustainable decisions that benefit urban trees and protect urban infrastructure. For example, a map of tree types in urban areas could assist in the diversification of urban forests to prevent tree pests and diseases, such as emerald ash borer, from devastating urban tree canopies [3, 4, 14, 27]. Tree maps could also help electrical companies monitor trees located near power lines. Trees falling onto power lines have caused many power outages and forest fires in recent years [1, 8]. In 2018, trees close to a powerline were the cause of the Camp Fire in California, a wildfire that killed 88 people [7, 28]. Some types of trees have lower resistance to wind than others, making them a hazard to surrounding infrastructure [28]. Knowing where different types of trees are located in relation to power lines can help cities better manage risks to public safety through tree trimming and removal efforts.

This problem is challenging for several reasons. Manual field surveys are time consuming and labor intensive. Therefore, they are not feasible for tree type identification over a large geographic extent. There is also limited dataset availability. Although extremely high spatial resolution LiDAR (i.e., centimeter resolution) and hyperspectral (i.e., sub-meter resolution) data have been used in some studies, these data do not cover expansive geographic regions [28]. Additionally, there is a lack of labeled training data. Thus, it is hard for machine learning approaches to be applied to solve this problem. Most previous literature has not made an effort to address this issue. Because of the large variation among trees of the same type, a large labeled training dataset is an essential aspect of solving this problem. Additionally, the process machine learning and deep learning methods depend on to produce their output is hard to interpret, so it can be difficult to improve model performance [9, 27].

There have been some successes in geospatial image detection using deep learning [10, 20, 25, 26, 30]. However, this method has been applied to detect objects that are clearly distinguishable by the human eye. In satellite imagery, trees appear to be green circles, and not much distinction is visible on the tree type level [24]. Other studies have classified trees based on hyperspectral imagery [5, 13, 29, 31], but these data are not available over a large geographic extent at a sufficiently high spatial resolution to be useful in widespread individual tree type classification. Likewise, high spatial resolution LiDAR data (i.e., centimeter resolution) has been used to study tree architecture [2, 12], but these datasets are also limited to custom studies. The most widely available remote sensing dataset with sufficient spatial resolution for studying individual trees remains RGB aerial imagery. Aerial imagery is collected in many urban counties, such as Hennepin and Ramsey Counties in Minnesota, at a 4–8 cm resolution, which is considered high spatial resolution [6, 16].

Our previous work has suggested that shadows could be used to classify trees based on their type, with some promising preliminary results using deep learning to classify trees in three guesses [28]. However, the actual differences in signatures of the shadows have not been characterized and their effectivity has not been quantified.

To overcome limitations of related work, our approach decomposes this problem down into two steps. First, a high-quality dataset of labeled tree shadows was created. Then, visual signatures of nine different types of trees in the dataset—ash, elm, hackberry, linden, locust, maple, oak, pine, and spruce—were analyzed and summarized. These observations were used by two human raters to test the feasibility of tree type identification based on shadows.

This work makes three major contributions. First, it offers a method to ensure the consistency of spatial shadow data. To create the dataset, shadows were annotated in aerial imagery. The tree type identification field survey used to create the dataset in this work was also verified manually to ensure its accuracy. Second, this work creates a labeled dataset of 4,613 trees based on the aforementioned methods to improve consistency between human raters. Third, this work details unique and distinct characteristics of different tree types based on shadow-only identification.

**Relevance to Big Data.** The aerial imagery used in this study had a very high spatial resolution (i.e., 4–8 cm) and required a large amount of storage space. For example, Ramsey County 7.6 cm resolution imagery covering 440 square kilometers was 12.8 GB. By this estimation, the continental United States, which spans approximately 8 million square kilometers, could require around 0.23 petabytes of storage. Furthermore, due to the large volume of trees in urban environments, an inventory of individual tree types that covers a large geographic region would be very big. The spatial and temporal resolution of aerial imagery will continue to improve in the future, making this problem even more relevant to big data forums in the coming years.

**Scope.** This paper focuses on the creation of a consistent and accurate dataset of tree shadows that can be used to train and test machine learning and deep learning models in the future. The following topics are outside the scope of this paper: analysis of extremely high spatial resolution hyperspectral and LiDAR data for tree type identification and application of machine learning and deep learning methods. Classification

here is limited to tree type classification by human raters, which is an important preliminary step to understanding the feasibility of machine learning-based approaches.

**Outline.** Section Two formally defines the problem. Section Three describes the approach we followed to create the labeled dataset and identify tree signatures. Section Four discusses the results of each phase of the approach. Section Five concludes the paper and suggests next steps of this study.

## 2 Problem Definition

### *Input*

- High spatial resolution RGB aerial imagery (i.e., 4–8 cm resolution).
- Low spatial resolution LiDAR data (i.e., 1 m resolution).
- Field survey point datasets covering limited regions of Hennepin County and Ramsey County (Minnesota, USA).
- Pre-existing domain knowledge (i.e., information from published field guides).

### *Output*

- Detailed rules and guidelines for classifying shadows by type.

### *Objective*

- Maximize tree type identification accuracy.

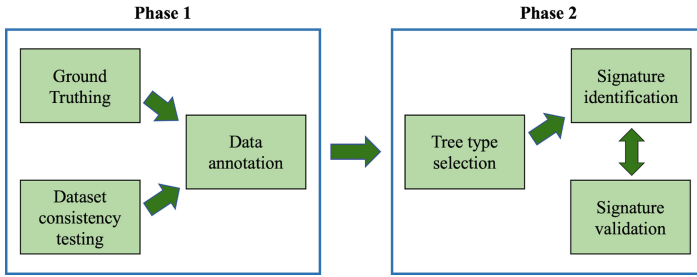
### *Constraints*

- Date and collection time of input aerial imagery was not available.
- Some datasets (i.e., high spatial resolution hyperspectral and LiDAR data) considered important in previous research aiming to identify trees was not available to the authors for the study region (as of September 23, 2019).
- Dataset of labeled tree shadows must be representative of the entire population.

## 3 Methods

The problem was broken down into two sub-problems. To determine important characteristics for distinguishing among nine different types of trees, the raters had to observe many samples of different types of trees. Thus, the first phase was creating a dataset with a high accuracy of labeled tree types and a high rate of consistency across shadow annotations. This phase consisted of ground truthing, dataset consistency testing, and data annotation (Fig. 2), described in further detail later in Sect. 3.1.

The goal of the second phase was to determine important characteristics for identifying tree types from shadows and to test the accuracy of identification of tree types based on their observed signatures. During phase two, certain tree types were chosen for closer observation based on their presence in the dataset: then signatures of these tree types' shadows were observed and validated (Fig. 2). This is described in further detail in Sect. 3.2.



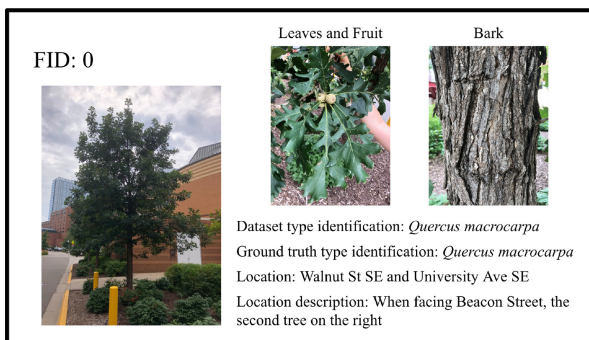
**Fig. 2.** Workflow of this study.

### 3.1 Labeled Dataset Creation

In the first phase of this study, pre-existing field survey datasets used to label the type of individual trees seen in aerial imagery were ground truthed (verified manually) to ensure accurate genus-level identifications. Dataset consistency testing was conducted to ensure consistent annotations between raters and guidelines for consistency testing were established. Using dataset consistency testing guidelines, a dataset of 4,613 shadows was annotated, and each shadow was labeled based on the verified field survey.

**Ground Truthing.** To assess the accuracy of the tree type identification dataset, two raters visited the locations of 100 tree points and identified them independently from the field survey. For safety reasons, the trees were selected from a 0.39 square kilometer area of the East Bank on the University of Minnesota campus. All unobstructed tree shadows within the test area were annotated. The 271 total annotated data points became the sample space for this experiment.

In ArcGIS, the genus and species data were removed from the feature class's attribute table and 100 of the 271 points were randomly selected. Each selected tree was visited, and its leaves, bark, and fruit, if present, were photographed. Figure 3 displays data collected for tree 0. Corresponding data were collected for every tree visited.



**Fig. 3.** Ground truthed data point.

After photographing the trees, raters identified their type using the following two field guides: *Minnesota Trees* and *The Sibley Guide to Trees* [19, 21]. The tree genus and species identifications from the original dataset were then matched with the raters' identifications. Percent agreement calculations were used to evaluate dataset accuracy based on ground truthed identifications.

**Dataset Consistency Testing.** To optimize consistency in data annotation between raters, a two-part dataset consistency test was designed to measure both binary agreement and spatial overlap. A fishnet, or grid covering the point dataset overlaid on the aerial imagery, was created and used as a coordinate system to randomly select data points for sampling. Grid extents were standardized. Two random numbers, representing the x and y coordinates of the grid, were generated. If a randomly generated cell in the fishnet contained points, all points within the cell were moved to the base of the corresponding tree trunk. Points were deleted if there were no local trees to which they could be assigned. Random cells were generated until the Hennepin 2015 dataset contained approximately 40 points, the Ramsey County 2015 dataset contained approximately 30 points, and the Ramsey County 2017 dataset contained approximately 30 points.

Each rater recorded a binary measure (1 or 0) for each data point indicating whether the tree was satisfactory for data annotation based on the clarity of the tree's branch structure. If determined suitable, the shadow was outlined according to guidelines established by the raters. Examples of annotations collected by each rater are shown in Fig. 4. Original annotations were traced for image clarity.



**Fig. 4.** Examples of shadow annotations collected by each rater during dataset consistency testing.

Once each rater had completed the dataset consistency test, binary agreement was assessed using the Cohen's Kappa Statistic [23, 32]. Cohen's Kappa is a statistic that accounts for chance when measuring agreement between raters on a scale of 0–1 where 0 is the probability of agreement based on random chance and 1 is perfect agreement. Equation (1) is the Cohen's Kappa statistic [23].  $P_o$  is the agreement between the two raters and  $p_e$  is the chance agreement expected to happen for the given experiment. Additionally, spatial agreement between tree shadow annotations was measured through a calculation of percent overlap between the outlines of each rater (Eq. (2)).

$$\kappa = \frac{po - pe}{1 - pe} \quad (1)$$

$$\frac{\sum \text{areas of intersection}}{\sum \text{areas of intersection} + \sum (\text{area of annotations} - \text{intersection area})} \times 100\% \quad (2)$$

Multiple phases of analysis were conducted until the Cohen's Kappa value met or exceeded a value of 0.800 and the average percent overlap over all three datasets met or exceeded a value of 75%.

After each phase of analysis, the raters reviewed and adjusted the codebook to increase their accuracy and consistency. For example, raters decided to annotate shadows as close to the tree as possible, and trees without distinct branches or canopies were no longer included in the dataset.

A similar procedure was employed to confirm that each rater's annotations remained consistent over time. The raters annotated randomly selected points and ensured that the agreement between raters met preset thresholds. Then, no less than two weeks later, each rater re-annotated the randomly selected points. Each rater's self-agreement was evaluated using the same methods as inter-rater consistency testing. If binary or spatial agreement was not satisfactory (i.e., less than 0.800 and 75%, respectively), the process was repeated.

**Annotation Methods.** Shadow annotation consisted of two phases. First, annotations were auto-generated for each point in the field survey dataset. Then, each annotation was reviewed and revised to ensure its quality and adherence to guidelines established in the dataset consistency phase of this research.

To auto-generate annotations for points in the field survey dataset, normalized height models (NHMs) were created in ArcGIS using digital elevation models and digital spatial models, both of which are LiDAR derivatives. The aerial imagery dataset used in this study was integrated, so the dates and times of collection were not available. Thus, shadow size and direction information had to be approximated for each NHM used. To do this, five shadows were randomly selected from the extent of each NHM. The quality of the shadow and clarity of the tree's canopy in the NHM were reviewed. If the height and width of the shadow were distinct and the range of the tree's canopy in the NHM was clear and relatively even, then the shadow was annotated. Otherwise, a new point was randomly selected from the field survey. Based on the shadow annotations and the canopies in the NHM, the angle of the shadow (azimuth), the height-of-tree to length-of-shadow ratio, and the width-of-shadow to height-of-tree ratio were calculated. The angles and ratios were averaged across the five randomly selected trees, and the extent of the NHM was noted. This information and the field survey dataset were used as inputs into a MATLAB script, which generated an initial set of shadow annotations.

Each annotation was then reviewed by the raters. Based on the dataset consistency guidelines developed earlier in this work, annotations corresponding to shadows with significant obstruction by other objects (i.e., cars, houses, other shadows), low contrast with their background, or a significant number of branches too narrow to be visible in the aerial imagery (i.e., the tree is very young) were deleted. Annotations

corresponding to trees removed before the aerial imagery was collected were also deleted. The size of each pre-generated annotation was checked and edited if necessary to ensure that it adequately suited the shape and extent of the tree’s shadow. The completed dataset contained 4,613 shadows.

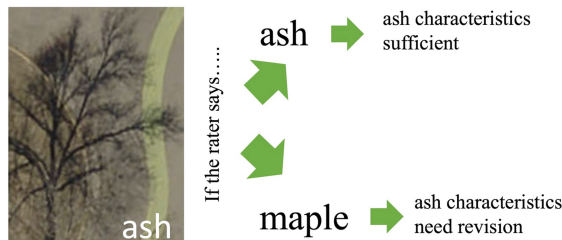
### 3.2 Identifying Tree Shadow Signatures in Aerial Imagery

To determine the differences among shadows of different types of trees, each tree type was observed in aerial imagery. Distinct characteristics were noted and, where possible, field guides with limited information about the characteristics of the silhouettes of different types of trees were reviewed to refine the observations and to guide further exploration.

**Tree Type Selection.** Nine types of trees were selected for identification in this study: ash, maple, linden, elm, oak, hackberry, locust, pine, and spruce. A tenth category included negative examples. The types of trees included in this analysis were the most common tree types in the initial 1,500 shadow dataset that the raters annotated, which included shadows from Hennepin County 2015 and Ramsey County 2015 data. Tree types were selected for identification based on their presence in the original dataset so that each category had an adequate number of samples.

**Visual Signature Identification Methods.** Raters reviewed samples from a small subsection of the dataset and created guidelines for distinguishing the nine chosen types of trees. To improve their ability to identify trees, raters predicted what type of tree they were annotating while growing the dataset and checked the accuracy of their guess with the field survey dataset. If they noticed other patterns when they were expanding the dataset, they added those patterns to the list of rules. Field guides were reviewed in order to gain additional insight into potential structural differences among the nine chosen types of trees.

**Feasibility Testing.** Feasibility testing was conducted to verify the observations made in the second phase of this study. In this phase of the study, two human raters identified tree types based on shadows. As they conducted feasibility testing, they refined tree type characteristics based on their performance (Fig. 5).



**Fig. 5.** Refinement based on feasibility testing.



*Ten-Tree-Type Feasibility Testing.* Each rater began by annotating a sample of shadows from the same aerial imagery dataset (Ramsey County 2018 7.6 cm resolution data). Information about the annotations was extracted to a separate file, where a portion of samples from each type of tree was randomly selected so that all ten categories of trees (maple, ash, pine, spruce, hackberry, locust, oak, linden, elm, and miscellaneous) were evenly represented. However, the distribution of tree types in the miscellaneous category was not even. From the evenly distributed dataset, 20 shadows were randomly selected, and a screenshot of each shadow was then taken at a consistent zoom (1:275). The screenshot and its identifying information were then put into a presentation, which became a test for the other rater. Each rater observed the 20 shadows that the other rater had annotated and noted distinct characteristics about the shadows, their top two identifications for the type of tree, and their final identification for the type of tree. This information was then compared to the identification included in the field survey point dataset, and the accuracy and Cohen's Kappa of the rater's classifications were calculated.

*Ash-Tree Feasibility Testing.* Methods followed for ash-tree feasibility testing were similar to methods followed in the ten-tree-type feasibility test except that there were only two categories, Ash and Miscellaneous.

## 4 Results

### 4.1 Experimental Goals

There were three goals of this research:

1. Create a substantial, accurate, and consistent dataset that can be used as a basis for machine learning and deep learning models in future research.
2. Observe and detail characteristics that are distinct among the shadows (profile geometries) of nine different types of trees and could be used to identify types of trees from their shadows in aerial imagery.
3. Test the feasibility of shadow-only-based tree type identification by human raters to validate the reliability of the characteristics observed in the second phase of this study and to better understand the potential for machine learning and deep learning classifiers to accurately classify trees by their shadows.

### 4.2 Labeled Dataset

To create an accurate dataset, ground truthing methods were employed to determine the accuracy level of the field surveys used in this study. Table 1 displays the agreement between the tree type identification point dataset and the tree genera and species data collected by the raters. The accuracy of the overall type level agreement was extremely high—97.9% (Table 1). One outlier is the accuracy of the elm type identification, which was very low. However, this can be discounted due to the small number of elm

trees in the sample space. This shows that the dataset labels were sufficiently accurate to observe structural trends among different types of trees. These results also show that there are inaccuracies, although small, in manual identification.

**Table 1.** Tree type identification dataset accuracy assessment.

Genus	Genus (type) agreement (%)	Species agreement (%)	Total number of trees
Ash	100.	0.00	1
Elm	50.0	50.0	2
Hackberry	100.	100.	1
Linden	100.	75.0	8
Locust	100.	100.	18
Maple	100.	76.2	21
Oak	100.	80.0	15
Pine	100.	85.7	7
Spruce	100.	100.	7
Miscellaneous	92.9	64.0	14
Inaccessible trees	–	–	6
Total	<b>97.9</b>	<b>83.3</b>	<b>100</b>

Dataset consistency methods were conducted to ensure that shadow annotations had consistent qualities and dimensions in relation to the size of the tree. First, these methods were utilized to measure the agreement between annotations created by each rater. The two raters met the preset consistency thresholds (based on averages) in their third phase of testing. The results for the third phase of testing are shown in Table 2.

**Table 2.** Results from third phase of dataset consistency testing.

Metric	Hennepin County 2015	Ramsey County 2015	Ramsey County 2017	Total
Cohen's Kappa	0.867	0.878	0.742	0.873
Percent agreement	93.5	93.9	90.9	93.7
Percent overlap	78.9	73.7	77.8	76.3

Testing was also conducted to measure the consistency of a rater's annotations over time. To do this, each rater annotated a set of 100 points that had met the agreement standards between raters. Two weeks later, the rater re-annotated the same subsection of the dataset and compared their annotations to the annotations they had collected two weeks prior using the same statistics mentioned previously. Both raters met the required standards for self-agreement.

## 5 Visual Signature Identification and Validation

**Signature Identification.** Observations made about each of the nine types of trees analyzed in this research are shown in Table 3, which displays an example, icon, and description of the nine types of trees in this study. The grey area in the icon represents “fuzziness”, while the black lines represent distinct branches. Icons are meant to highlight certain characteristics described in the narrative and are not to scale. The comments and examples included in the table were the most representative of the trees contained in the dataset.

Authors made several observations throughout this analysis. Previous research has noted that maple trees and ash trees are notable for their opposite branching. However, distinguishing whether a tree had opposite or alternate branching based on its shadow in aerial imagery was challenging and led to many incorrect tree type classifications. Additionally, young trees look different from their fully matured counterparts [11, 12]. As a result, the appearance of younger trees is not included in Table 3. One of the most distinct characteristics among tree types was “fuzziness” of the tree. Fuzziness is seen in aerial imagery because many small branches are not clearly visible due to the resolution. However, this characteristic is especially noteworthy in ash trees, which have a lot of fuzziness around their terminal branches.

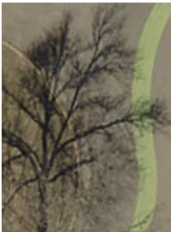

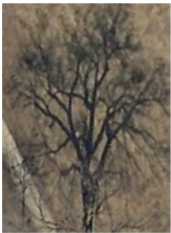

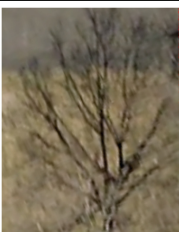

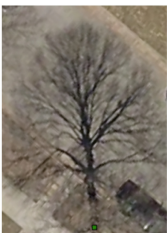
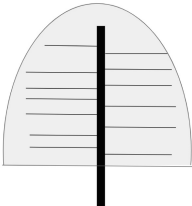
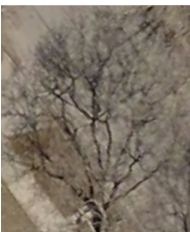
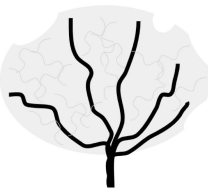
**Feasibility Testing.** The purpose of feasibility testing was to verify whether the signatures observed earlier in this study were valid. Two types of tests were conducted: ten-tree-type feasibility tests and ash-tree feasibility tests. The results of the ten-tree-type feasibility tests are shown in Figs. 6 and 7.


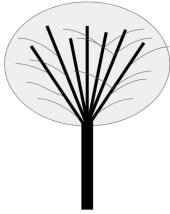
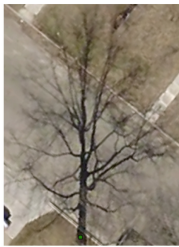
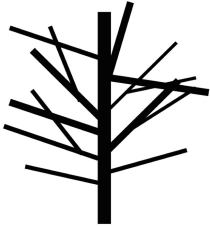




Over the course of ten-tree-type feasibility testing, each rater improved their accuracy in one and two type identification predictions. During trials 1 and 2, raters strictly followed the guidelines they had created. However, after viewing many more shadow samples for each type of tree and noticing heavy variation among shadows of the same tree type, raters attempted a more “instinct-based” approach, which likely accounted for the dip in accuracy in the third trial. In trial 4, raters used a tree type narrative they had created. The narrative included descriptions and example shadows for each type of tree. Comparing the tree shadows they were observing to other tree shadows in the dataset with known labels helped each rater improve their identification accuracy, as shown in the figure. Accuracy improved when comparing unknown samples to known samples, which demonstrates the potential for machine learning techniques to be applied to this problem.

Additionally, the identification of the 100 trees during the ground truth phase of this study took 25 h to complete. On the other hand, the identification of 20 trees during the feasibility testing phase of this study took approximately one hour to complete. Thus, if this method was scaled up, trees could potentially be classified up to five times faster than traditional methods, even before the application of machine learning methods.

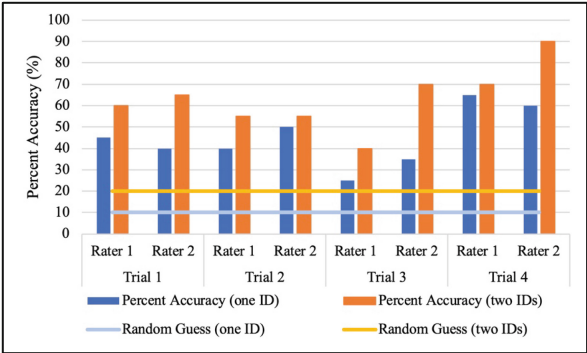
Results for ash-tree feasibility testing are shown in Figs. 8 and 9. In Fig. 8, the blue columns represent identification accuracies. The horizontal blue line represents the percent accuracy achieved based on a random guess. Results improved after the first trial and demonstrate potential for identification of ash trees from their shadows. In Fig. 9, blue columns represent Cohen’s Kappa values. The Cohen’s Kappa values for

**Table 3.** Tree structure observations.

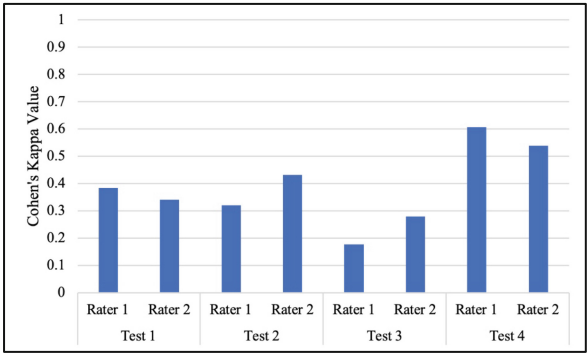
Type	Example	Icon	Distinctive Characteristics in Shadows
Ash			<ul style="list-style-type: none"> <li>• Trunk splits at start of canopy [17]</li> <li>• Fuzzy, feather-like ends of distinct terminal branches</li> <li>• Uneven canopy shape surrounding the main branches</li> </ul>
Elm			<ul style="list-style-type: none"> <li>• Trunk splits into many distinct branches [17]</li> <li>• Wavy branches</li> <li>• Light fuzziness</li> <li>• Branches grow at steep upward angle, but may angle down at ends (water fountain shape) [17]</li> </ul>
Hackberry			<ul style="list-style-type: none"> <li>• Distinct U-shaped fork in trunk occurs near the initial branches of the tree canopy [21]</li> <li>• Straighter branches than locust trees, which resemble hackberry trees</li> <li>• Many distinct branches create smooth canopy outline</li> </ul>
Linden			<ul style="list-style-type: none"> <li>• Very round, smooth, oval canopy, especially at top [21]</li> <li>• Relatively flat branch angle.</li> <li>• Trunk usually doesn't split, but if it does, the split is narrow</li> <li>• Few dark, distinct branches</li> <li>• High degree of uniform fuzziness</li> </ul>
Locust			<ul style="list-style-type: none"> <li>• Wide trunk split, sometimes U-shaped</li> <li>• Many distinct branches</li> <li>• Squiggly, spirally fuzziness</li> <li>• Distinct branches are squiggly and may have a few significantly curved gnarls [17, 21]</li> </ul>

Maple			<ul style="list-style-type: none"> <li>• Distinct rounded or champagne glass-shaped canopy [17, 21]</li> <li>• Minimal branching below the trunk split</li> <li>• Many straight distinct branches at a steep upward angle [17, 21] and smooth uniform fuzziness [21]</li> <li>• Narrow branches are close to each other [21]</li> </ul>
Oak			<p>Possibility 1 (shown in icon):</p> <ul style="list-style-type: none"> <li>• Rarely have trunk split</li> <li>• Relatively flat branches skinny compared to trunk [21]</li> <li>• Haphazard branch arrangement</li> </ul> <p>Possibility 2:</p> <ul style="list-style-type: none"> <li>• Dark, gnarled branches and trunk [17]</li> </ul>
Pine			<ul style="list-style-type: none"> <li>• Mostly black shadows</li> <li>• Speckled [17]</li> <li>• Rounded top</li> <li>• Short compared to width</li> </ul>
Spruce			<ul style="list-style-type: none"> <li>• Dense black shadows</li> <li>• Jagged edges</li> <li>• Pointy top [17]</li> <li>• Skinny compared to height [17]</li> </ul>

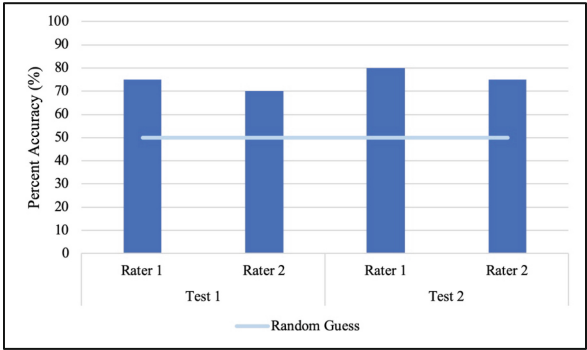
ash-tree-feasibility testing are more consistent than the ten-tree-type feasibility test Cohen's Kappa value, suggesting that the differences are easier to observe. These clear differences may make ash classification more conducive to machine learning.



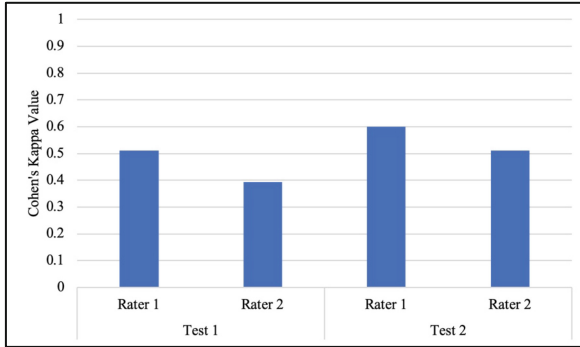
**Fig. 6.** Percent accuracy for each rater for ten-tree-type feasibility testing.



**Fig. 7.** Cohen's Kappa value achieved by each rater for ten-tree-type feasibility testing.



**Fig. 8.** Percent accuracy by each rater for ash-tree feasibility testing.



**Fig. 9.** Cohen's Kappa value achieved by each rater for ash-tree feasibility testing. (Color figure online)

## 6 Conclusions and Future Work

In this study, we created a dataset of 4,613 clear and distinct labeled tree shadows, established rules and guidelines for identifying trees based solely on shadows, and refined these rules to maximize feasibility accuracy scores. Our study shows there are distinct differences among the profile geometries of trees that allow them to be accurately classified by human raters faster than traditional field survey methods. Distinct differences among tree types also demonstrates the feasibility of classifying trees based on their shadows using automated methods.

The domain-knowledge outlined in this paper can be used to create machine learning and deep learning models that identify types of individual trees in aerial imagery covering large regions. Due to the large amount of high spatial resolution aerial imagery as well as the large volume of individual trees, supercomputers or cloud computers may be necessary to improve the scalability of this approach [18]. Additionally, as higher resolution hyperspectral and LiDAR data become more widely available, these methods could be explored in tandem to create more reliable classification results.

**Acknowledgements.** This study is supported by the US NSF under Grants No. 1901099, 1737633, 1541876, 1029711, IIS-1320580, 0940818 and IIS-1218168, the USDOD under Grants HM0210-13-1-0005, USDA under Grant No. 2017-51181-27222, ARPA-E under Grant No. DE-AR0000795, NIH under Grant No. UL1 TR002494, KL2 TR002492 and TL1 TR0024-93, and the OVPR U-Spatial and Minnesota Supercomputing Institute at the University of Minnesota. Aerial imagery used in this work and shown in this paper was supplied by Hennepin and Ramsey County. Field surveys were provided by the City of St. Paul Forestry Unit and the University of Minnesota. We thank Kim Koffolt for improving the readability of this paper. We also thank the Spatial Computing Group for their feedback throughout this work.

## References

1. Acevedo, N.: Single fallen tree on power line leaves 900K without power. NBC News. Associated Press, Puerto Rico (2018). <https://www.nbcnews.com/storyline/puerto-rico-crisis/puerto-rico-fallen-tree-power-line-leaves-900k-without-power-n865506>. Accessed 1 Feb 2019
2. Åkerbloma, M., Raumonena, P., Mäkipääb, R., Kaasalainenaa, M.: Automatic tree species recognition with quantitative structure models. *Remote Sens. Environ.* **191**, 1–12 (2017). <https://doi.org/10.1016/j.rse.2016.12.002>
3. BenDor, T., Metcalf, S., Fontenot, L., Sangunett, B., Hannon, B.: Modeling the spread of the Emerald Ash Borer. *Int. J. Ecol. Model. Syst. Ecol.* **197**(1), 221–236 (2006). <https://doi.org/10.1016/j.ecolmodel.2006.03.003>
4. Citywide EAB management strategies. <https://www.stpaul.gov/departments/parks-recreation/natural-resources/forestry/disease-pest-management/citywide-eab>. Accessed 9 Nov 2018
5. Dalponte, M., Ørka, H.O., Gobakken, T., Gianelle, D., Næsset, E.: Tree species classification in boreal forests with hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **51**(5), 2632–2645 (2012). <https://doi.org/10.1109/TGRS.2012.2216272>
6. Elevation and Imagery. <https://gis-hennepin.opendata.arcgis.com/pages/imagery>. Accessed 25 Aug 2018
7. Fuller, T.: Three Weeks After Fire, Official Search for Dead is Completed. *The New York Times*. <https://www.nytimes.com/2018/11/29/us/victims-california-fires-missing.html>. Accessed 1 Feb 2019
8. Gold, R., Blunt, K., Smith, R.: PG&E sparked at least 1,500 California fires: now the utility faces collapse. *Wall Street J.* (2019). <https://www.wsj.com/articles/pg-e-sparked-at-least-1-500-california-fires-now-the-utility-faces-collapse-11547410768>. Accessed 1 Feb 2019
9. Hutson, M.: Has artificial intelligence become alchemy? *Science* **360**(6388), 478 (2018). <https://doi.org/10.1126/science.360.6388.478>
10. Lee, S.H., Chan, C.S., Wilkin, P., Remagnino, P.: Deep-plant: plant identification with convolutional neural networks. In: 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, Canada, 27–30 September 2015. <https://doi.org/10.1109/ICIP.2015.7350839>
11. Leonardi, C., Stagi, F.: *The Architecture of Trees*. Princeton Architectural Press, Hudson (2019). <https://doi.org/10.1109/ICIP.2015.7350839>
12. Malhi, Y., et al.: New perspectives on the ecology of tree structure and tree communities through terrestrial laser scanning. *Interface Focus*. **8**(2), 1–10 (2018). <https://doi.org/10.1098/rsfs.2017.0052>
13. Murfitt, J., He, Y., Yang, J., Mui, A., De Mille, K.: Ash decline assessment in emerald ash borer infested natural forests using high spatial resolution images. *Remote Sens.* **8**(3), 256 (2016). <https://doi.org/10.3390/rs8030256>
14. Nisley, R.G.: Emerald ash borer research: a decade of progress on an expanding pest problem. *Northern Res. Stat. Res. Rev.* **20**, 1–5 (2013)
15. Nowak, D.J., Dwyer, J.F.: Understanding the benefits and costs of urban forest ecosystems. In: Kuser, J.E. (ed.) *Urban and Community Forestry in the Northeast*, pp. 25–46. Springer, Dordrecht (2007). [https://doi.org/10.1007/978-1-4020-4289-8\\_2](https://doi.org/10.1007/978-1-4020-4289-8_2)
16. Orthophotos/Aerial 2015 (MapServer). <https://maps.co.ramsey.mn.us/arcgis/rest/services/OrthoPhotos/Aerial2015/MapServer>. Accessed 25 Aug 2018
17. Petrides, G.A., Wehr, J.: *A Field Guide to Eastern Trees: Eastern United States and Canada, Including the Midwest*. Houghton Mifflin Company, New York (1998)



18. Prasad, S.K., et al.: Parallel processing over spatial-temporal datasets from geo, bio, climate and social science communities: a research roadmap. In: 2017 IEEE International Congress on Big Data (BigData Congress), Honolulu, HI, USA, 25–30 June 2017. <https://doi.org/10.1109/BigDataCongress.2017.39>
19. Rathke, D.M.: Minnesota Trees. University of Minnesota Extension Service, Minnesota (2006)
20. Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, 21–26 July 2017. <https://doi.org/10.1109/CVPR.2017.690>
21. Sibley, D.A.: The Sibley Guide to Trees. Alfred A. Knopf Inc., New York (2009)
22. Tyrväinen, L., Pauleit, S., Seeland, K., de Vries, S.: Benefits and uses of urban forests and trees. In: Konijnendijk, C.C., Nilsson, K., Randrup, T.B., Schipperijn, J. (eds.) *Urban Forests and Trees*, pp. 81–144. Springer, New York (2005). [https://doi.org/10.1007/3-540-27684-X\\_5](https://doi.org/10.1007/3-540-27684-X_5)
23. Viera, A.J., Garrett, J.M.: Understanding interobserver agreement: the kappa statistic. *Family Med.* **37**(5), 360–363 (2005)
24. Xie, Y., Bao, H., Shekhar, S., Knight, J.: A TIMBER framework for mining urban tree inventories using remote sensing datasets. In: Proceedings of 2018 IEEE International Conference on Data Mining (ICDM), Singapore, 17–20 November 2018. <https://doi.org/10.1109/ICDM.2018.00183>
25. Xie, Y., Bhojwani, R., Shekhar, S., Knight, J.: An unsupervised augmentation framework for deep learning based geospatial object detection: a summary of results. In: Proceedings of the ACM SIGSPATIAL International Conference on Advancements in Geographic Information Systems (SIGSPATIAL 2018), Seattle, WA, 6–9 November 2018. <https://doi.org/10.1145/3274895.3274901>
26. Xie, Y., Cai, J., Bhojwani, R., Shekhar, S., Knight, J.: A locally constrained YOLO framework for detecting small and densely distributed building footprints. *Int. J. Geograph. Inf. Sci.* 1–25 (2019). <https://doi.org/10.1080/13658816.2019.1624761>
27. Xie, Y., Gupta, J., Li, Y., Shekhar, S.: Transforming smart cities with spatial computing. In: 2018 IEEE International Smart Cities Conference (ISC2), Kansas City, MO, USA, 16–19 September 2018. <https://doi.org/10.1109/ISC2.2018.8656800>
28. Xie, Y., Shekhar, S., Feiock, R., Knight, J.: Revolutionizing tree management via intelligent spatial techniques. Paper presented at the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (SIGSPATIAL 2019), Chicago, IL, USA, 5–8 November 2019. <https://doi.org/10.1145/3347146.3359066>
29. Yu, C., Li, M., Zhang, M.: Classification of dominant tree species in an urban forest park using the remote sensing image of WorldView-2. In: 8th International Congress on Image and Signal Processing (CISP), pp. 742–747 (2015). <https://doi.org/10.1109/CISP.2015.7407976>
30. Yuan, J.: Learning building extraction in aerial scenes with convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **40**(11), 2793–2798 (2017). <https://doi.org/10.1109/TPAMI.2017.2750680>
31. Zhang, C., Qiu, F.: Mapping individual tree species in an urban forest using airborne lidar data and hyperspectral imagery. *Photogram. Eng. Remote Sens.* **78**(10), 1079–1087 (2012). <https://doi.org/10.14358/PERS.78.10.1079>
32. Freelon, D.: ReCal2: Reliability for 2 Coders. <https://dfreelon.org/utlis/recalfont/recal2/>. Accessed 21 Sep 2018